

APRIL/MAY 2024

**23PCS21 — DATA MINING AND
WAREHOUSING**

Time : Three hours

Maximum : 75 marks

SECTION A — (10 × 2 = 20 marks)

Answer ALL questions.

1. What are the basic data mining tasks?
2. List out the key issues in data mining.
3. Define Classification.
4. What are rule-based algorithms in classification?
5. What are association rules and why are they used in data mining?
6. How are outliers handled in clustering algorithms?
7. What is a snowflake schema and how does it differ from a star schema?
8. How are OLAP tools integrated with the internet?
9. Why is it important to build a data warehouse?
10. What are some organization issues in data warehousing?

SECTION B — ($5 \times 5 = 25$ marks)

Answer ALL questions.

11. (a) Explain the difference between data mining and knowledge discovery in databases.

Or

- (b) Discuss the Social implications of data mining.

12. (a) Compare Statistical based and distance based algorithms in classification.

Or

- (b) What are the combining techniques used in data mining algorithms?

13. (a) Define Clustering and its importance in data analysis.

Or

- (b) Compare hierarchical and partitional clustering algorithms.

14. (a) Elucidate the concept of data marts and their role in data warehousing.

Or

- (b) List out the characteristics of data warehousing.

15. (a) Discuss the architectural strategies involved in building a data warehousing.

Or

- (b) Discuss the applications of data warehousing in government sectors.

SECTION C — ($3 \times 10 = 30$ marks)

Answer any THREE questions.

16. Analyze the Statistical perspective on data mining and its significance.

17. Evaluate the strengths and weakness of neural network based algorithms in data mining.

18. Explain the significance incremental rules in association rule mining.

19. Discuss the role of OLAP tools in leveraging data for strategic decision making.

20. Explore the design considerations for data content in a data warehousing.